

SEQUENCE RECORD

(1) GENERAL INDICATIONS:

(i) APPLICANT:

- (A) NAME: Deutsches Krebsforschungszentrum
- (B) STREET: Im Neuenheimer Feld 280
- (C) TOWN: Heidelberg
- (E) COUNTRY: Germany
- (F) POSTAL CODE: 69120

(ii) TITLE OF THE INVENTION: Multivalent Antibody Constructs

(iii) NUMBER OF SEQUENCES: 17

(iv) COMPUTER-READABLE VERSION:

- (A) DATA CARRIER: floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, version #1.30 (EPA)

(2) INDICATIONS AS TO SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1698 base pairs
- (B) KIND: nucleotide
- (C) STRAND TYPE: single strand
- (D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: genome DNA

(iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) POSITION: 28..1689

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) POSITION: 28..1689

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTCATTA AAGAGGAGAA ATTAACC ATG AAA TAC CTA TTG CCT ACG GCA
Met Lys Tyr Leu Leu Pro Thr Ala
1 5

GCC Ala	GCT Ala	GGC Gly	TTG Leu	CTG Leu	CTG Leu	CTG Leu	GCA Ala	GCT Ala	CAG Gln	CCG Pro	GCC Ala	ATG Met	GCG Ala	CAG Gln	GTG Val	99
101520																
CAA Gln	CTG Leu	CAG Gln	CAG Gln	TCT Ser	GGG Gly	GCT Ala	GAA Glu	CTG Leu	GCA Ala	AGA Arg	CCT Pro	GGG Gly	GCC Ala	TCA Ser	GTG Val	147
25303540																
AAG Lys	ATG Met	TCC Ser	TGC Cys	AAG Lys	GCT Ala	TCT Ser	GGC Gly	TAC Tyr	ACC Thr	TTT Phe	ACT Thr	AGG Arg	TAC Tyr	ACG Thr	ATG Met	195
455055																
CAC His	TGG Trp	GTA Val	AAA Lys	CAG Gln	AGG Arg	CCT Pro	GGA Gly	CAG Gln	GGT Gly	CTG Leu	GAA Glu	TGG Trp	ATT Ile	GGA Gly	TAC Tyr	243
606570																
ATT Ile	AAT Asn	CCT Pro	AGC Ser	CGT Arg	GGT Gly	TAT Tyr	ACT Thr	AAT Asn	TAC Tyr	AAT Asn	CAG Gln	AAG Lys	TTC Phe	AAG Lys	GAC Asp	291
758085																
AAG Lys	GCC Ala	ACA Thr	TTG Leu	ACT Thr	ACA Thr	GAC Asp	AAA Lys	TCC Ser	TCC Ser	AGC Ser	ACA Thr	GCC Ala	TAC Tyr	ATG Met	CAA Gln	339
9095100																
CTG Leu	AGC Ser	AGC Ser	CTG Leu	ACA Thr	TCT Ser	GAG Glu	GAC Asp	TCT Ser	GCA Ala	GTC Val	TAT Tyr	TAC Tyr	TGT Cys	GCA Ala	AGA Arg	387
105110115																
TAT Tyr	TAT Tyr	GAT Asp	GAT Asp	CAT His	TAC Tyr	AGC Ser	CTT Leu	GAC Asp	TAC Tyr	TGG Trp	GGC Gly	CAA Gln	GGC Gly	ACC Thr	ACT Thr	435
125130135																
CTC Leu	ACA Thr	GTC Val	TCC Ser	TCA Ser	GCC Ala	AAA Lys	ACA Thr	ACA Thr	CCC Pro	AAG Lys	CTT Leu	GGC Gly	GGT Gly	GAT Asp	ATC Ile	483
140145150																
TTG Leu	CTC Leu	ACC Thr	CAA Gln	ACT Thr	CCA Pro	GCT Ala	TCT Ser	TTG Leu	GCT Ala	GTG Val	TCT Ser	CTA Leu	GGG Gly	CAG Gln	AGG Arg	531
155160165																
GCC Ala	ACC Thr	ATC Ile	TCC Ser	TGC Cys	AAG Lys	GCC Ala	AGC Ser	CAA Gln	AGT Ser	GTT Val	GAT Asp	TAT Tyr	GAT Asp	GGT Gly	GAT Asp	579
170175180																
AGT Ser	TAT Tyr	TTG Leu	AAC Asn	TGG Trp	TAC Tyr	CAA Gln	CAG Gln	ATT Ile	CCA Pro	GGA Gly	CAG Gln	CCA Pro	CCC Pro	AAA Lys	CTC Leu	627
185190195200																
CTC Leu	ATC Ile	TAT Tyr	GAT Asp	GCA Ala	TCC Ser	AAT Asn	CTA Leu	GTT Val	TCT Ser	GGG Gly	ATC Ile	CCA Pro	CCC Pro	AGG Arg	TTT Phe	675
205210215																
AGT Ser	GGC Gly	AGT Ser	GGG Gly	TCT Ser	GGG Gly	ACA Thr	GAC Asp	TTC Phe	ACC Thr	CTC Leu	AAC Asn	ATC Ile	CAT His	CCT Pro	GTG Val	723
220225230																

GAG Glu	AAG Lys	GTG Val 235	GAT Asp	GCT Ala	GCA Ala	ACC Thr	TAT Tyr 240	CAC His	TGT Cys	CAG Gln	CAA Gln	AGT Ser 245	ACT Thr	GAG Glu	GAT Asp	771
CCG Pro	TGG Trp 250	ACG Thr	TTC Phe	GGT Gly	GGA Gly	GGC Gly 255	ACC Thr	AAG Lys	CTG Leu	GAA Glu	ATC Ile 260	AAA Lys	CGG Arg	GCT Ala	GAT Asp	819
GCT Ala 265	GCG Ala	GCC Ala	GCT Ala	GGT Gly	GGT Gly 270	GGT Gly	GGT Gly	TCT Ser	GGC Gly	GGC Gly 275	GGT Gly	GGT Gly	AGC Ser	GGT Gly	GGT Gly 280	867
GGC Gly	GGC Gly	TCC Ser	GGT Gly	GGT Gly 285	GGT Gly	GGT Gly	AGC Ser	CAG Gln	GTG Val 290	CAG Gln	CTG Leu	CAG Gln	CAG Gln	TCT Ser 295	GGG Gly	915
GCT Ala	GAG Glu	CTG Leu 300	GTG Val	AGG Arg	CCT Pro	GGG Gly	TCC Ser	TCA Ser 305	GTG Val	AAG Lys	ATT Ile	TCC Ser	TGC Cys 310	AAG Lys	GCT Ala	963
TCT Ser	GGC Gly	TAT Tyr 315	GCA Ala	TTC Phe	AGT Ser	AGC Ser	TAC Tyr 320	TGG Trp	ATG Met	AAC Asn	TGG Trp	GTG Val 325	AAG Lys	CAG Gln	AGG Arg	1011
CCT Pro	GGA Gly 330	CAG Gln	GGT Gly	CTT Leu	GAG Glu	TGG Trp 335	ATT Ile	GGA Gly	CAG Gln	ATT Ile	TGG Trp 340	CCT Pro	GGA Gly	GAT Asp	GGT Gly	1059
GAT Asp 345	ACT Thr	AAC Asn	TAC Tyr	AAT Asn	GGA Gly 350	AAG Lys	TTC Phe	AAG Lys	GGT Gly	AAA Lys 355	GCC Ala	ACT Thr	CTG Leu	ACT Thr	GCA Ala 360	1107
GAC Asp	GAA Glu	TCC Ser	TCC Ser	AGC Ser 365	ACA Thr	GCC Ala	TAC Tyr	ATG Met	CAA Gln 370	CTC Leu	AGC Ser	AGC Ser	CTA Leu	GCA Ala 375	TCT Ser	1155
GAG Glu	GAC Asp	TCT Ser	GCG Ala 380	GTC Val	TAT Tyr	TTC Phe	TGT Cys	GCA Ala 385	AGA Arg	CGG Arg	GAG Glu	ACT Thr	ACG Thr 390	ACG Thr	GTA Val	1203
GGC Gly	CGT Arg	TAT Tyr 395	TAC Tyr	TAT Tyr	GCT Ala	ATG Met	GAC Asp 400	TAC Tyr	TGG Trp	GGT Gly	CAA Gln	GGA Gly 405	ACC Thr	TCA Ser	GTC Val	1251
ACC Thr	GTC Val 410	TCC Ser	TCA Ser	GCC Ala	AAA Lys	ACA Thr 415	ACA Thr	CCC Pro	AAG Lys	CTT Leu	GGC Gly 420	GGT Gly	GAT Asp	ATC Ile	GTG Val	1299
CTC Leu 425	ACT Thr	CAG Gln	TCT Ser	CCA Pro	GCA Ala 430	ATC Ile	ATG Met	TCT Ser	GCA Ala	TCT Ser 435	CCA Pro	GGG Gly	GAG Glu	AAG Lys	GTC Val 440	1347
ACC Thr	ATG Met	ACC Thr	TGC Cys	AGT Ser 445	GCC Ala	AGC Ser	TCA Ser	AGT Ser	GTA Val 450	AGT Ser	TAC Tyr	ATG Met	AAC Asn	TGG Trp 455	TAC Tyr	1395

CAG CAG AAG TCA GGC ACC TCC CCC AAA AGA TGG ATT TAT GAC ACA TCC	1443
Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg Trp Ile Tyr Asp Thr Ser	
460 465 470	
AAA CTG GCT TCT GGA GTC CCT GCT CAC TTC AGG GGC AGT GGG TCT GGG	1491
Lys Leu Ala Ser Gly Val Pro Ala His Phe Arg Gly Ser Gly Ser Gly	
475 480 485	
ACC TCT TAC TCT CTC ACA ATC AGC GGC ATG GAG GCT GAA GAT GCT GCC	1539
Thr Ser Tyr Ser Leu Thr Ile Ser Gly Met Glu Ala Glu Asp Ala Ala	
490 495 500	
ACT TAT TAC TGC CAG CAG TGG AGT AGT AAC CCA TTC ACG TTC GGC TCG	1587
Thr Tyr Tyr Cys Gln Gln Trp Ser Ser Asn Pro Phe Thr Phe Gly Ser	
505 510 515 520	
GGG ACA AAG TTG GAA ATA AAC CGG GCT GAT ACT GCA CCA ACT GGA TCC	1635
Gly Thr Lys Leu Glu Ile Asn Arg Ala Asp Thr Ala Pro Thr Gly Ser	
525 530 535	
GAA CAA AAG CTG ATC TCA GAA GAA GAC CTA AAC TCA CAT CAC CAT CAC	1683
Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ser His His His His	
540 545 550	
CAT CAC TAATCTAGA	1698
His His	

(2) INDICATIONS AS TO ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 554 amino acids

(B) KIND: amino acid

(D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala	
1 5 10 15	
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Ala Glu	
20 25 30	
Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly	
35 40 45	
Tyr Thr Phe Thr Arg Tyr Thr Met His Trp Val Lys Gln Arg Pro Gly	
50 55 60	
Gln Gly Leu Glu Trp Ile Gly Tyr Ile Asn Pro Ser Arg Gly Tyr Thr	
65 70 75 80	

Asn	Tyr	Asn	Gln	Lys 85	Phe	Lys	Asp	Lys	Ala	Thr	Leu	Thr	Thr	Asp	Lys 95
Ser	Ser	Ser	Thr 100	Ala	Tyr	Met	Gln	Leu 105	Ser	Ser	Leu	Thr	Ser 110	Glu	Asp
Ser	Ala	Val 115	Tyr	Tyr	Cys	Ala	Arg 120	Tyr	Tyr	Asp	Asp	His 125	Tyr	Ser	Leu
Asp	Tyr 130	Trp	Gly	Gln	Gly	Thr 135	Thr	Leu	Thr	Val	Ser 140	Ser	Ala	Lys	Thr
Thr 145	Pro	Lys	Leu	Gly	Gly 150	Asp	Ile	Leu	Leu	Thr 155	Gln	Thr	Pro	Ala	Ser 160
Leu	Ala	Val	Ser	Leu 165	Gly	Gln	Arg	Ala	Thr 170	Ile	Ser	Cys	Lys	Ala 175	Ser
Gln	Ser	Val	Asp 180	Tyr	Asp	Gly	Asp	Ser 185	Tyr	Leu	Asn	Trp	Tyr 190	Gln	Gln
Ile	Pro	Gly 195	Gln	Pro	Pro	Lys	Leu 200	Leu	Ile	Tyr	Asp	Ala 205	Ser	Asn	Leu
Val	Ser 210	Gly	Ile	Pro	Pro	Arg 215	Phe	Ser	Gly	Ser	Gly 220	Ser	Gly	Thr	Asp
Phe 225	Thr	Leu	Asn	Ile	His 230	Pro	Val	Glu	Lys	Val 235	Asp	Ala	Ala	Thr	Tyr 240
His	Cys	Gln	Gln	Ser 245	Thr	Glu	Asp	Pro	Trp 250	Thr	Phe	Gly	Gly	Gly 255	Thr
Lys	Leu	Glu	Ile 260	Lys	Arg	Ala	Asp	Ala 265	Ala	Ala	Ala	Gly	Gly 270	Gly	Gly
Ser	Gly	Gly 275	Gly	Gly	Ser	Gly	Gly 280	Gly	Gly	Ser	Gly	Gly 285	Gly	Gly	Ser
Gln	Val 290	Gln	Leu	Gln	Gln	Ser 295	Gly	Ala	Glu	Leu	Val 300	Arg	Pro	Gly	Ser
Ser 305	Val	Lys	Ile	Ser	Cys 310	Lys	Ala	Ser	Gly	Tyr 315	Ala	Phe	Ser	Ser	Tyr 320
Trp	Met	Asn	Trp	Val 325	Lys	Gln	Arg	Pro	Gly 330	Gln	Gly	Leu	Glu	Trp 335	Ile
Gly	Gln	Ile	Trp 340	Pro	Gly	Asp	Gly	Asp 345	Thr	Asn	Tyr	Asn	Gly 350	Lys	Phe
Lys	Gly	Lys 355	Ala	Thr	Leu	Thr	Ala 360	Asp	Glu	Ser	Ser	Ser 365	Thr	Ala	Tyr

Met Gln Leu Ser Ser Leu Ala Ser Glu Asp Ser Ala Val Tyr Phe Cys
 370 375 380

Ala Arg Arg Glu Thr Thr Thr Val Gly Arg Tyr Tyr Tyr Ala Met Asp
 385 390 395 400

Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Lys Thr Thr
 405 410 415

Pro Lys Leu Gly Gly Asp Ile Val Leu Thr Gln Ser Pro Ala Ile Met
 420 425 430

Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Ser Ala Ser Ser
 435 440 445

Ser Val Ser Tyr Met Asn Trp Tyr Gln Gln Lys Ser Gly Thr Ser Pro
 450 455 460

Lys Arg Trp Ile Tyr Asp Thr Ser Lys Leu Ala Ser Gly Val Pro Ala
 465 470 475 480

His Phe Arg Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser
 485 490 495

Gly Met Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser
 500 505 510

Ser Asn Pro Phe Thr Phe Gly Ser Gly Thr Lys Leu Glu Ile Asn Arg
 515 520 525

Ala Asp Thr Ala Pro Thr Gly Ser Glu Gln Lys Leu Ile Ser Glu Glu
 530 535 540

Asp Leu Asn Ser His His His His His His
 545 550

(2) INDICATIONS AS TO ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1653 base pairs
- (B) KIND: nucleotide
- (C) STRAND TYPE: single strand
- (D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: genome DNA

(iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) POSITION: 28..1644

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GCC ACC ATC TCC TGC AAG GCC AGC CAA AGT GTT GAT TAT GAT GGT GAT 579
Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val Asp Tyr Asp Gly Asp
170 175 180

AGT Ser 185	TAT Tyr	TTG Leu	AAC Asn	TGG Trp	TAC Tyr 190	CAA Gln	CAG Gln	ATT Ile	CCA Pro	GGA Gly 195	CAG Gln	CCA Pro	CCC Pro	AAA Lys	CTC Leu 200	627
CTC Leu	ATC Ile	TAT Tyr	GAT Asp	GCA Ala 205	TCC Ser	AAT Asn	CTA Leu	GTT Val	TCT Ser 210	GGG Gly	ATC Ile	CCA Pro	CCC Pro	AGG Arg 215	TTT Phe	675
AGT Ser	GGC Gly	AGT Ser	GGG Gly 220	TCT Ser	GGG Gly	ACA Thr	GAC Asp	TTC Phe 225	ACC Thr	CTC Leu	AAC Asn	ATC Ile	CAT His 230	CCT Pro	GTG Val	723
GAG Glu	AAG Lys	GTG Val 235	GAT Asp	GCT Ala	GCA Ala	ACC Thr	TAT Tyr 240	CAC His	TGT Cys	CAG Gln	CAA Gln	AGT Ser 245	ACT Thr	GAG Glu	GAT Asp	771
CCG Pro	TGG Trp 250	ACG Thr	TTC Phe	GGT Gly	GGA Gly	GGC Gly 255	ACC Thr	AAG Lys	CTG Leu	GAA Glu	ATC Ile 260	AAA Lys	CGG Arg	GCT Ala	GAT Asp	819
GCT Ala 265	GCG Ala	GCC Ala	GCT Ala	GGT Gly	GGC Gly 270	CCA Pro	GGG Gly	TCG Ser	CAG Gln	GTG Val 275	CAG Gln	CTG Leu	CAG Gln	CAG Gln	TCT Ser 280	867
GGG Gly	GCT Ala	GAG Glu	CTG Leu	GTG Val 285	AGG Arg	CCT Pro	GGG Gly	TCC Ser	TCA Ser 290	GTG Val	AAG Lys	ATT Ile	TCC Ser	TGC Cys 295	AAG Lys	915
GCT Ala	TCT Ser	GGC Gly	TAT Tyr 300	GCA Ala	TTC Phe	AGT Ser	AGC Ser	TAC Tyr 305	TGG Trp	ATG Met	AAC Asn	TGG Trp 310	GTG Val	AAG Lys	CAG Gln	963
AGG Arg	CCT Pro	GGA Gly 315	CAG Gln	GGT Gly	CTT Leu	GAG Glu	TGG Trp 320	ATT Ile	GGA Gly	CAG Gln	ATT Ile	TGG Trp 325	CCT Pro	GGA Gly	GAT Asp	1011
GGT Gly	GAT Asp 330	ACT Thr	AAC Asn	TAC Tyr	AAT Asn	GGA Gly 335	AAG Lys	TTC Phe	AAG Lys	GGT Gly	AAA Lys 340	GCC Ala	ACT Thr	CTG Leu	ACT Thr	1059
GCA Ala 345	GAC Asp	GAA Glu	TCC Ser	TCC Ser	AGC Ser 350	ACA Thr	GCC Ala	TAC Tyr	ATG Met	CAA Gln 355	CTC Leu	AGC Ser	AGC Ser	CTA Leu	GCA Ala 360	1107
TCT Ser	GAG Glu	GAC Asp	TCT Ser	GCG Ala 365	GTC Val	TAT Tyr	TTC Phe	TGT Cys	GCA Ala 370	AGA Arg	CGG Arg	GAG Glu	ACT Thr	ACG Thr 375	ACG Thr	1155
GTA Val	GGC Gly	CGT Arg	TAT Tyr 380	TAC Tyr	TAT Tyr	GCT Ala	ATG Met	GAC Asp 385	TAC Tyr	TGG Trp	GGT Gly	CAA Gln	GGA Gly 390	ACC Thr	TCA Ser	1203
GTC Val	ACC Thr	GTC Val 395	TCC Ser	TCA Ser	GCC Ala	AAA Lys	ACA Thr 400	ACA Thr	CCC Pro	AAG Lys	CTT Leu 405	GGC Gly	GGT Gly	GAT Asp	ATC Ile	1251

GTG	CTC	ACT	CAG	TCT	CCA	GCA	ATC	ATG	TCT	GCA	TCT	CCA	GGG	GAG	AAG	1299
Val	Leu	Thr	Gln	Ser	Pro	Ala	Ile	Met	Ser	Ala	Ser	Pro	Gly	Glu	Lys	
	410					415					420					
GTC	ACC	ATG	ACC	TGC	AGT	GCC	AGC	TCA	AGT	GTA	AGT	TAC	ATG	AAC	TGG	1347
Val	Thr	Met	Thr	Cys	Ser	Ala	Ser	Ser	Ser	Val	Ser	Tyr	Met	Asn	Trp	
425					430					435					440	
TAC	CAG	CAG	AAG	TCA	GGC	ACC	TCC	CCC	AAA	AGA	TGG	ATT	TAT	GAC	ACA	1395
Tyr	Gln	Gln	Lys	Ser	Gly	Thr	Ser	Pro	Lys	Arg	Trp	Ile	Tyr	Asp	Thr	
			445						450					455		
TCC	AAA	CTG	GCT	TCT	GGA	GTC	CCT	GCT	CAC	TTC	AGG	GGC	AGT	GGG	TCT	1443
Ser	Lys	Leu	Ala	Ser	Gly	Val	Pro	Ala	His	Phe	Arg	Gly	Ser	Gly	Ser	
		460						465					470			
GGG	ACC	TCT	TAC	TCT	CTC	ACA	ATC	AGC	GGC	ATG	GAG	GCT	GAA	GAT	GCT	1491
Gly	Thr	Ser	Tyr	Ser	Leu	Thr	Ile	Ser	Gly	Met	Glu	Ala	Glu	Asp	Ala	
	475						480					485				
GCC	ACT	TAT	TAC	TGC	CAG	CAG	TGG	AGT	AGT	AAC	CCA	TTC	ACG	TTC	GGC	1539
Ala	Thr	Tyr	Tyr	Cys	Gln	Gln	Trp	Ser	Ser	Asn	Pro	Phe	Thr	Phe	Gly	
	490					495					500					
TCG	GGG	ACA	AAG	TTG	GAA	ATA	AAC	CGG	GCT	GAT	ACT	GCA	CCA	ACT	GGA	1587
Ser	Gly	Thr	Lys	Leu	Glu	Ile	Asn	Arg	Ala	Asp	Thr	Ala	Pro	Thr	Gly	
505					510					515					520	
TCC	GAA	CAA	AAG	CTG	ATC	TCA	GAA	GAA	GAC	CTA	AAC	TCA	CAT	CAC	CAT	1635
Ser	Glu	Gln	Lys	Leu	Ile	Ser	Glu	Glu	Asp	Leu	Asn	Ser	His	His	His	
			525						530					535		
CAC	CAT	CAC	TAATCTAGA													1653
His	His	His														

(2) INDICATIONS AS TO ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 539 amino acids

(B) KIND: amino acid

(D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Lys	Tyr	Leu	Leu	Pro	Thr	Ala	Ala	Ala	Gly	Leu	Leu	Leu	Leu	Ala	
1				5					10					15		
Ala	Gln	Pro	Ala	Met	Ala	Gln	Val	Gln	Leu	Gln	Gln	Ser	Gly	Ala	Glu	
		20					25						30			
Leu	Ala	Arg	Pro	Gly	Ala	Ser	Val	Lys	Met	Ser	Cys	Lys	Ala	Ser	Gly	
		35					40					45				

Tyr 50	Phe	Thr	Arg	Tyr	Thr 55	Met	His	Trp	Val	Lys 60	Gln	Arg	Pro	Gly	
Gln 65	Gly	Leu	Glu	Trp	Ile 70	Gly	Tyr	Ile	Asn	Pro 75	Ser	Arg	Gly	Tyr	Thr 80
Asn	Tyr	Asn	Gln	Lys 85	Phe	Lys	Asp	Lys	Ala 90	Thr	Leu	Thr	Thr	Asp 95	Lys
Ser	Ser	Ser	Thr 100	Ala	Tyr	Met	Gln	Leu 105	Ser	Ser	Leu	Thr	Ser 110	Glu	Asp
Ser	Ala	Val 115	Tyr	Tyr	Cys	Ala	Arg 120	Tyr	Tyr	Asp	Asp	His 125	Tyr	Ser	Leu
Asp	Tyr 130	Trp	Gly	Gln	Gly	Thr 135	Thr	Leu	Thr	Val	Ser 140	Ser	Ala	Lys	Thr
Thr 145	Pro	Lys	Leu	Gly	Gly 150	Asp	Ile	Leu	Leu	Thr 155	Gln	Thr	Pro	Ala	Ser 160
Leu	Ala	Val	Ser	Leu 165	Gly	Gln	Arg	Ala	Thr 170	Ile	Ser	Cys	Lys	Ala 175	Ser
Gln	Ser	Val	Asp 180	Tyr	Asp	Gly	Asp	Ser 185	Tyr	Leu	Asn	Trp	Tyr 190	Gln	Gln
Ile	Pro	Gly 195	Gln	Pro	Pro	Lys	Leu 200	Leu	Ile	Tyr	Asp	Ala 205	Ser	Asn	Leu
Val	Ser 210	Gly	Ile	Pro	Pro	Arg 215	Phe	Ser	Gly	Ser	Gly 220	Ser	Gly	Thr	Asp
Phe 225	Thr	Leu	Asn	Ile	His 230	Pro	Val	Glu	Lys	Val 235	Asp	Ala	Ala	Thr	Tyr 240
His	Cys	Gln	Gln	Ser 245	Thr	Glu	Asp	Pro	Trp 250	Thr	Phe	Gly	Gly	Gly 255	Thr
Lys	Leu	Glu	Ile 260	Lys	Arg	Ala	Asp	Ala 265	Ala	Ala	Ala	Gly	Gly 270	Pro	Gly
Ser	Gln	Val 275	Gln	Leu	Gln	Gln	Ser 280	Gly	Ala	Glu	Leu	Val 285	Arg	Pro	Gly
Ser	Ser 290	Val	Lys	Ile	Ser	Cys 295	Lys	Ala	Ser	Gly	Tyr 300	Ala	Phe	Ser	Ser
Tyr 305	Trp	Met	Asn	Trp	Val 310	Lys	Gln	Arg	Pro	Gly 315	Gln	Gly	Leu	Glu	Trp 320
Ile	Gly	Gln	Ile	Trp 325	Pro	Gly	Asp	Gly	Asp 330	Thr	Asn	Tyr	Asn	Gly 335	Lys

Phe Lys Gly Lys Ala Thr Leu Thr Ala Asp Glu Ser Ser Ser Thr Ala
 340 345 350
 Tyr Met Gln Leu Ser Ser Leu Ala Ser Glu Asp Ser Ala Val Tyr Phe
 355 360 365
 Cys Ala Arg Arg Glu Thr Thr Thr Val Gly Arg Tyr Tyr Tyr Ala Met
 370 375 380
 Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Lys Thr
 385 390 395 400
 Thr Pro Lys Leu Gly Gly Asp Ile Val Leu Thr Gln Ser Pro Ala Ile
 405 410 415
 Met Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Ser Ala Ser
 420 425 430
 Ser Ser Val Ser Tyr Met Asn Trp Tyr Gln Gln Lys Ser Gly Thr Ser
 435 440 445
 Pro Lys Arg Trp Ile Tyr Asp Thr Ser Lys Leu Ala Ser Gly Val Pro
 450 455 460
 Ala His Phe Arg Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
 465 470 475 480
 Ser Gly Met Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
 485 490 495
 Ser Ser Asn Pro Phe Thr Phe Gly Ser Gly Thr Lys Leu Glu Ile Asn
 500 505 510
 Arg Ala Asp Thr Ala Pro Thr Gly Ser Glu Gln Lys Leu Ile Ser Glu
 515 520 525
 Glu Asp Leu Asn Ser His His His His His His
 530 535

- (2) INDICATIONS AS TO ID NO: 5:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 57 base pairs
 - (B) KIND: nucleotide
 - (C) STRAND TYPE: single strand
 - (D) TOPOLOGY: linear
 - (ii) KIND OF MOLECULE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
 - (iii) HYPOTHETICAL: no
 - (iv) ANTISENSE: no
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

57

- ```
(ii) KIND OF MOLECULE: other nucleic acid
 (A) DESCRIPTION: /desc = "primer"
(iii) HYPOTHETICAL: no
(iv) ANTISENSE: no
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
```

45

- 34

## (2) INDICATIONS AS TO ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 348 base pairs
- (B) KIND: nucleotide
- (C) STRAND TYPE: single strand
- (D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: genome DNA

(iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) POSITION: 1..348

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) POSITION: 1..348

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA TCC TCC | 48  |
| Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser |     |
| 1 5 10 15                                                       |     |
| GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA | 96  |
| Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln |     |
| 20 25 30                                                        |     |
| ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC | 144 |
| Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe |     |
| 35 40 45                                                        |     |
| GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG | 192 |
| Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu |     |
| 50 55 60                                                        |     |
| TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA | 240 |
| Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val |     |
| 65 70 75 80                                                     |     |
| TCT CTC GAG AAA AGA GAG GCT GAA GCT GAA TTC CAG GTG CAA CTG CAG | 288 |
| Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Gln Val Gln Leu Gln |     |
| 85 90 95                                                        |     |
| CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG AAG ATG TCC | 336 |
| Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser |     |
| 100 105 110                                                     |     |
| TGC AAG GCT TCT                                                 | 348 |
| Cys Lys Ala Ser                                                 |     |
| 115                                                             |     |

## 2) INDICATIONS AS TO ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 amino acids
- (B) KIND: amino acid
- (D) TOPOLOGY: linear

## (ii) KIND OF MOLECULE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1 5 10 15
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
 20 25 30
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
 35 40 45
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 50 55 60
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65 70 75 80
Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Gln Val Gln Leu Gln
 85 90 95
Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser
 100 105 110
Cys Lys Ala Ser
 115

```

## (2) INDICATIONS AS TO ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 354 base pairs
- (B) KIND: nucleotide
- (C) STRAND TYPE: single strand
- (D) TOPOLOGY: linear

## (ii) KIND OF MOLECULE: genome DNA

## (iii) HYPOTHETICAL: no

## (iv) ANTISENSE: no

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) POSITION: 1..354

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) POSITION: 1..354

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | AGA | TTT | CCT | TCA | ATT | TTT | ACT | GCT | GTT | TTA | TTC | GCA | GCA | TCC | TCC | 48  |
| Met | Arg | Phe | Pro | Ser | Ile | Phe | Thr | Ala | Val | Leu | Phe | Ala | Ala | Ser | Ser |     |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GCA | TTA | GCT | GCT | CCA | GTC | AAC | ACT | ACA | ACA | GAA | GAT | GAA | ACG | GCA | CAA | 96  |
| Ala | Leu | Ala | Ala | Pro | Val | Asn | Thr | Thr | Thr | Glu | Asp | Glu | Thr | Ala | Gln |     |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ATT | CCG | GCT | GAA | GCT | GTC | ATC | GGT | TAC | TCA | GAT | TTA | GAA | GGG | GAT | TTC | 144 |
| Ile | Pro | Ala | Glu | Ala | Val | Ile | Gly | Tyr | Ser | Asp | Leu | Glu | Gly | Asp | Phe |     |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GAT | GTT | GCT | GTT | TTG | CCA | TTT | TCC | AAC | AGC | ACA | AAT | AAC | GGG | TTA | TTG | 192 |
| Asp | Val | Ala | Val | Leu | Pro | Phe | Ser | Asn | Ser | Thr | Asn | Asn | Gly | Leu | Leu |     |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| TTT | ATA | AAT | ACT | ACT | ATT | GCC | AGC | ATT | GCT | GCT | AAA | GAA | GAA | GGG | GTA | 240 |
| Phe | Ile | Asn | Thr | Thr | Ile | Ala | Ser | Ile | Ala | Ala | Lys | Glu | Glu | Gly | Val |     |
| 65  |     |     |     |     | 70  |     |     |     | 75  |     |     |     |     |     | 80  |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| TCT | CTC | GAG | AAA | AGA | GAG | GCT | GAA | GCT | GAA | TTC | ATG | GCG | CAG | GTG | CAA | 288 |
| Ser | Leu | Glu | Lys | Arg | Glu | Ala | Glu | Ala | Glu | Phe | Met | Ala | Gln | Val | Gln |     |
|     |     |     |     | 85  |     |     |     | 90  |     |     |     |     |     | 95  |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| CTG | CAG | CAG | TCT | GGG | GCT | GAA | CTG | GCA | AGA | CCT | GGG | GCC | TCA | GTG | AAG | 336 |
| Leu | Gln | Gln | Ser | Gly | Ala | Glu | Leu | Ala | Arg | Pro | Gly | Ala | Ser | Val | Lys |     |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ATG | TCC | TGC | AAG | GCT | TCT |     |     |     |     |     |     |     |     |     |     | 354 |
| Met | Ser | Cys | Lys | Ala | Ser |     |     |     |     |     |     |     |     |     |     |     |
|     |     | 115 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

- 2) INDICATIONS AS TO ID NO:11:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 118 amino acids
    - (B) KIND: amino acid
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Arg | Phe | Pro | Ser | Ile | Phe | Thr | Ala | Val | Leu | Phe | Ala | Ala | Ser | Ser |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ala | Leu | Ala | Ala | Pro | Val | Asn | Thr | Thr | Thr | Glu | Asp | Glu | Thr | Ala | Gln |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ile | Pro | Ala | Glu | Ala | Val | Ile | Gly | Tyr | Ser | Asp | Leu | Glu | Gly | Asp | Phe |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu  
 50 55 60  
 Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val  
 65 70 75 80  
 Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Met Ala Gln Val Gln  
 85 90 95  
 Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys  
 100 105 110  
 Met Ser Cys Lys Ala Ser  
 115

- (2) INDICATIONS AS TO ID NO: 12:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 42 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TCACACAGAA TTCTTAGATC TATTAAAGAG GAGAAATTAA CC

42

- (2) INDICATIONS AS TO ID NO: 13:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 40 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:



AGCACACGAT ATCACCGCCA AGCTTGGGTG TTGTTTTGGC

40

- (2) INDICATIONS AS TO ID NO: 14:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 43 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14

AGCACACAAG CTTGGCGGTG ATATCTTGCT CACCCAACT CCA

43

- (2) INDICATIONS AS TO ID NO: 15:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 57 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"

- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15

AGCACACTCT AGAGACACAC AGATCTTTAG TGATGGTGAT GGTGATGTGA GTTTAGG

57

- (2) INDICATIONS AS TO ID NO: 16:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 33 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear

- (ii) KIND OF MOLECULE: other nucleic acid
  - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CAGCCGGCCA TGGCGCAGGT GCAACTGCAG CAG

33

- (2) INDICATIONS AS TO ID NO: 17:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 102 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

TATATACTGC AGCTGCACCT GGCTACCACC ACCACCGGAG CCGCCACCAC CGCTACCACC  
GCCGCCAGAA CCACCACCAC CAGCGGCCGC AGCATCAGCC CG

60

102

1012200 "1012200" 0012200